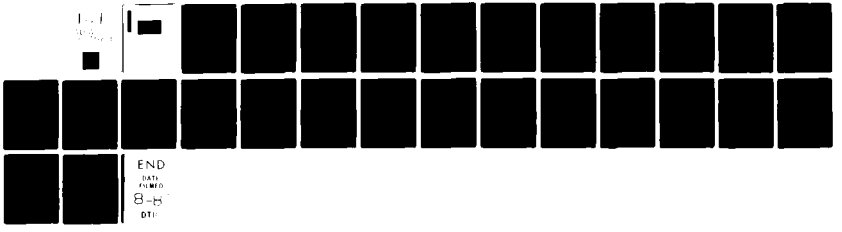
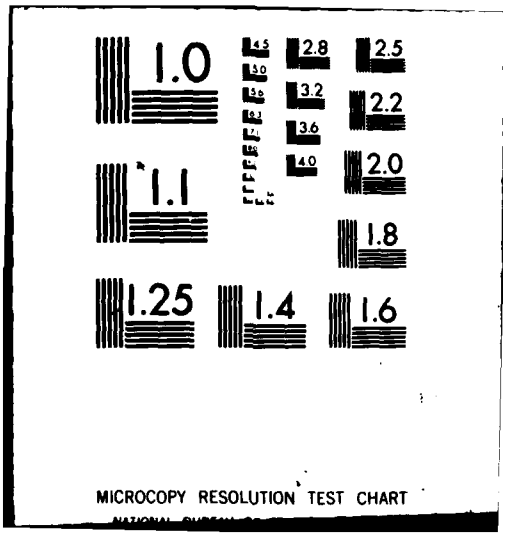
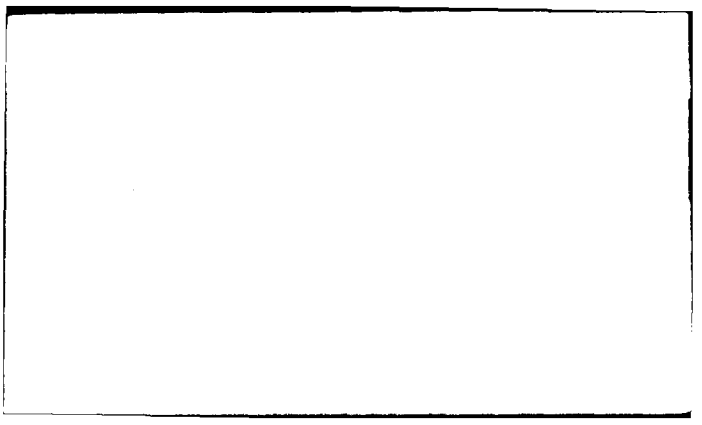
AD-A117 529 DESMATICS INC STATE COLLEGE PA F/G 9/2
FACTOR SCREENING IN SIMULATION: EVALUATION OF A RANDOM BALANCE/--ETC(U)
JUL 82 C A MAURO, D E SMITH N00014-79-C-0650
UNCLASSIFIED TR-113-6 NL



END
DATE FILMED
8-82
DTIC


1.0
1.1
1.25
1.4
1.6
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.8
MICROCOPY RESOLUTION TEST CHART



DESMATICS, INC.

P. O. Box 618
State College, Pa. 16801
Phone: (814) 238-9621

Applied Research in Statistics - Mathematics - Operations Research

FACTOR SCREENING IN SIMULATION: EVALUATION OF A RANDOM BALANCE/ PLACKETT-BURMAN STRATEGY

by

Carl A. Mauro
and
Dennis E. Smith

TECHNICAL REPORT NO. 113-6

July 1982

This study was supported by the Office of Naval Research under Contract No. N00014-79-C-0650, Task No. NR 277-291

TABLE OF CONTENTS

I. INTRODUCTION

Complex computer simulation models (computer codes) usually must be studied by performing a number of simulation runs. That is, we must "experiment" with the simulation model. Standard experimental designs, however, can very easily require more computer runs than are reasonable or affordable, especially when many factors (i.e., input variables) are present. In such cases it may be beneficial to invest a relatively small number of runs in a preliminary experiment aimed at determining which factors are the most important. By screening out those factors which appear to be relatively unimportant, we can concentrate the major experimental effort and expense on the important factors. The smaller the proportion of important effects, the more it is to our advantage to conduct a screening experiment.

Because running a large, complex simulation can be costly and time-consuming, the number of runs available for screening is generally severely limited. Typically, a supersaturated situation exists. That is, the number of factors to be studied exceeds the number of available screening runs. Although a number of screening strategies satisfying this constraint have been proposed (see, for example, Kleijnen [4], Srivastava [9], and Smith and Mauro [8]), no definite guidelines for selecting a screening strategy have been established, nor have the available methods been evaluated and compared systematically.

In this report we evaluate and discuss a two-stage screening procedure which we refer to as an RP strategy. The first stage of this strategy is a K factor random balance (RB) experiment where K denotes the total number of factors to be screened (see Satterthwaite [7] and Budne [1]); the second

stage, a follow-up to the first stage, is a Plackett-Burman (PB) experiment (see Plackett and Burman [6]). A factor is included in the second-stage experiment only if it is determined to have a significant effect in the first-stage experiment.

In a previous Desmatics, Inc. technical report, Mauro and Smith [5] formally investigated and reported the performance of the RP class of strategies in the case of zero error variance (i.e., where the simulation response is observed without random error). Our purpose here is to extend the performance analysis to the general case of nonzero error variance. As a statistical basis to assess this strategy, we assume the following model:

$$y_i = \beta_0 + \sum_{j=1}^{K} \beta_j x_{ij} + \varepsilon_i \qquad (1.1)$$

where:

(1) y_i is the value of the response(i.e., output variable) in the $i^{\underline{th}}$ simulation run,

(2) K is the total number of factors to be screened, each of which is at two levels coded +1 and -1,

(3) x_{ij} is +1 or -1 depending on the level of the $j^{\underline{th}}$ factor during the $i^{\underline{th}}$ simulation run,

(4) β_0 is a component common to all responses, and $\beta_j (j \geq 1)$ is the linear effect of the $j^{\underline{th}}$ factor,

and (5) the error terms, ε_i, are independent, have common distribution $N(0, \sigma^2)$, and are independent of every design variable (x_{ij}) in the model.

For detecting the factors having major effects, model (1.1) is gener-

ally sufficient. In essence, it is a first-order Taylor series approximation to an actual relationship between output and input variables. Performance evaluation will be restricted to this model.

The basic function of a screening strategy is to sort all the factors into essentially two groups: a group containing the so-called important factors and a group containing the so-called unimportant factors. The factors classified as important can then be subjected to more detailed study in subsequent experimentation. In an RP strategy, a factor is classified as important only if it reaches the second-stage experiment and is subsequently determined to have a significant effect. All other factors are classified as unimportant.

The actual importance (or unimportance) of a factor, however, will more generally depend on the magnitude of its effect relative to that of experimental error, σ, and to the magnitudes of the other effects present. From a practical standpoint, the greater (lesser) the degree of importance, the larger should be the probability of classifying the factor as important (unimportant). In this report, we summarize performance in terms of a strategy's sensitivity (i.e., power) for declaring a factor important and in terms of the expected number of runs the strategy requires.

II. PRELIMINARY DISCUSSION

In a two-level (±1) RB design, each column of the design matrix consists of N/2 +1's and N/2 -1's where N (an even number) denotes the total number of runs. The allocation of +1's and -1's to each column is made randomly so that all possible configurations of N/2 +1's and N/2 -1's (there are $C^{N}_{N/2}$ in all) are equally likely, with each column receiving an independent allocation. The condition that every column of the design matrix has an equal number of runs at the high (+1) and low (-1) levels assures us that estimates of the individual factor effects are unconfounded with the overall mean effect, which can be represented by a column of N +1's.

RB designs are attractive since they are easy to prepare for any N and K. Furthermore, unlike more orthodox designs where some mathematical relationship usually exists between N and K, we can select N independently of K in an RB design. This affords us a great deal of flexibility and simplification, particularly if K is large. Of course, a RB design also has its corresponding disadvantages. Indeed, the main objection to such designs is that they confound factors to a random degree and have no specialized or unique method of analysis. (For a more complete discussion of the pros and cons of random designs, the reader may consult Technometrics, Vol. 1, No. 2, May 1959.) The results of the present report may help to resolve some of the controversy surrounding random balance experimentation.

Because of the independent allocation of factor levels to each design column, practically any technique used to analyze data without random balance properties can be used to analyze any (sufficiently small) subset of factors in a random balance design. This is accomplished by ignoring any

factor not included in the particular subset being analyzed. The effects of the ignored factors, however, will be absorbed into the error component of the model. In the simplest case, we can consider each factor separately and apply some standard statistical analysis. In this report, we consider a standard F-test applied separately to each factor as the method of analysis for random balance data, and, for simplicity, conduct each F-test at the same level of significance, α_1.

Factors having a significant F-ratio in the first stage are carried over to the second stage to be tested in a PB follow-up experiment. Because PB designs are orthogonal, the second stage separates any confounding between the factors carried over from the RB first stage. Factors not formally included in the second-stage experiment should be held at a constant level so not to bias any of the second-stage estimates. The results of the second stage can be analyzed by the usual analysis of variance procedures for factorial experiments.

Although we can specify the number of first-stage runs N, the number of second-stage runs M will depend on the number of factors S carried over from the first stage. For reasons of economy, we generally employ the smallest PB design having at least S+1 runs. However, to avoid possible saturation and thus no degrees of freedom to estimate experimental error, the convention we will follow here is to employ the smallest PB design that guarantees at least one error degree of freedom. Since PB designs are only available for numbers of runs that are multiples of four, we can obtain a minimum of one and a maximum of four error degrees of freedom by following this procedure.

We can express M mathematically as $M = B(S+1)$ where

$$B(x) = x + 4 - x \pmod 4 \ . \qquad (2.1)$$

A useful approximation to B(x) is given by

$$B(x) \approx x + 2.5 \ . \tag{2.2}$$

The total number of runs R required by an RP strategy, therefore, will be $N+M = N+B(S+1)$. Note that since M is random, so is R. Using (2.2) we can approximate E(R) by

$$E(R) \approx N + E(S) + 3.5 \ . \tag{2.3}$$

Since $|B(x) - (x+2.5)| \leq 1.5$, the approximation in (2.3) can differ from E(R) by at most 1.5.

In summary, an RP strategy is completely determined by the number of first stage runs, N, and the significance levels of the first and second stage tests, α_1 and α_2, respectively. (We assume the same level of significance is used in all second-stage testing.) We denote such a strategy by $RP(N, \alpha_1, \alpha_2)$. The parameters N, α_1, and α_2 are at our disposal in selecting and specifying an RP screening strategy. In Section III we show how these quantities affect the performance of an RP screening plan.

III. PERFORMANCE EVALUATION

In this section we examine the performance of an RP(N, α_1, α_2) strategy. We discuss the first and second stages individually, followed by consideration of the combined stages. Of primary interest is the probability that a given factor is declared important by an RP strategy, or equivalently the probability that the factor tests significant in both the first and second stages. In general, this probability is too complex to be evaluated analytically. In lieu of an analytic solution, we develop an approximation to this probability. We also discuss an approximation to the expected value of R, the total number of runs required by an RP strategy.

A. FIRST STAGE: RANDOM BALANCE DESIGN

In matrix terms we can write model (1.1) as $\underline{y} = \beta_0\underline{1} + \underline{X}\underline{\beta} + \underline{\varepsilon}$ where $\underline{1}$ is an N x 1 vector of +1's, $\underline{y}$ is an N x 1 vector of responses, $\underline{\varepsilon}$ is an N x 1 vector of error terms, $\underline{\beta}$ is an K x 1 vector of factor effects, and $\underline{X}$ is an N x K design matrix. In a random balance experiment, $\underline{X} = [\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_K]$ is a stochastic matrix whose $j^{\underline{th}}$ column, $\underline{x}_j$, is an N x 1 vector consisting of a random arrangement of N/2 +1's and N/2 -1's. The K column vectors of $\underline{X}$ are stochastically independent.

The simple least squares estimator of $\beta_j (j \geq 1)$ is given by

$$\hat{\beta}_j = (\bar{y}_{+j} - \bar{y}_{-j})/2 \qquad (3.1)$$

where $\bar{y}_{+j}(\bar{y}_{-j})$ is the average value of the response over the N/2 runs at the +1 (-1) level of the $j^{\underline{th}}$ factor. (By simple least squares we mean that each β_j is estimated ignoring all other factors.) In matrix terms

$$\hat{\beta}_j = (\underline{x}_j'\underline{y})/N = [(\sum_{i=1}^{K} \beta_i \underline{x}_j'\underline{x}_i) + \underline{x}_j'\underline{\varepsilon}]/N \tag{3.2}$$

Thus,
$$E(\hat{\beta}_j) = (1/N)[\sum_{i=1}^{K} \beta_i E(\underline{x}_j'\underline{x}_i) + E(\underline{x}_j'\underline{\varepsilon})], \tag{3.3}$$

and
$$V(\hat{\beta}_j) = (1/N^2)[\sum_{i=1}^{K} \beta_i^2 V(\underline{x}_j'\underline{x}_i) + V(\underline{x}_j'\underline{\varepsilon})], \tag{3.4}$$

where in (3.4) we make use of the fact that $\underline{x}_j'\underline{\varepsilon}$, $\underline{x}_j'\underline{x}_1$, $\underline{x}_j'\underline{x}_2$, ..., $\underline{x}_j'\underline{x}_K$ are mutually independent for fixed j.

It is clear that the exact sampling distribution of $\hat{\beta}_j$ is intractable. In the Appendix, however, we show that

$$E(\hat{\beta}_j) = \beta_j, \tag{3.5}$$

$$V(\hat{\beta}_j) = (\tau^2 - \beta_j^2)/(N-1) + \sigma^2/N, \tag{3.6}$$

and
$$\text{cov}(\hat{\beta}_i, \hat{\beta}_j) = \beta_i\beta_j/(N-1), \tag{3.7}$$

where $\tau^2 = \sum_{m=1}^{K} \beta_m^2$. The simple least squares estimator defined in (3.1) is therefore an unbiased estimator of β_j, although its variance can be seriously inflated. Furthermore, the correlation between $\hat{\beta}_i$ and $\hat{\beta}_j$ is roughly (replacing N - 1 in (3.6) with N)

$$\text{corr}(\hat{\beta}_i, \hat{\beta}_j) \approx \beta_i\beta_j/V_iV_j \tag{3.8}$$

where $V_m = [\tau^2 + \sigma^2 - \beta_m^2]^{\frac{1}{2}}$.

The correlation in (3.8) is a measure of the confounding between $\hat{\beta}_i$ and $\hat{\beta}_j$. Notice that increasing N cannot decrease the confounding in an RB design where simple least squares is used. Moreover, the degree of confounding between $\hat{\beta}_i$ and $\hat{\beta}_j$ is dependent upon the magnitudes of other effects in

the model.

Because each factor is at two levels, the standard F-test to test $H_0:\beta_j = 0$ versus $H_1:\beta_j \neq 0$ is equivalent to a simple two-sample t-test between the high (+1) and the low (-1) levels of the $j^{\underline{th}}$ factor. The associated test statistic t_j is given by

$$t_j = \hat{\beta}_j / [SSE_j / N(N-2)]^{\frac{1}{2}} \tag{3.9}$$

where SSE_j is the familiar analysis of variance notation for the error sum of squares of factor j. Computationally,

$$SSE_j = \sum_H (y_i - \bar{y}_{+j})^2 + \sum_L (y_i - \bar{y}_{-j})^2 \tag{3.10}$$

where the first (second) summation is taken over the N/2 observations at the high (low) level of the $j^{\underline{th}}$ factor. Alternative computational formulas are

$$SSE_j = \sum_{i=1}^{N} y_i^2 - N\hat{\beta}_j^2 - N\bar{y}^2 \tag{3.11}$$

$$= \sum_{i=1}^{N} (y_i - \bar{y})^2 - N\hat{\beta}_j^2 \tag{3.12}$$

$$= \sum_{i=1}^{N} (y_i - \bar{y} - \hat{\beta}_j x_{ij})^2 \tag{3.13}$$

where $\bar{y}$ is the overall mean of the responses.

We reject H_0 in favor of H_1 if the observed value of $|t_j|$ equals or exceeds $t(N-2;\alpha_1/2)$, the upper $100(1-\alpha_1/2)$ percentage point of "Student's" t-distribution having (N - 2) degrees of freedom. We should point out, however, that t_j does not truly follow a t-distribution since the assumptions of so-called normal theory are not met exactly. As a result, the true size

of the test may differ from α_1. It is well known, though, that a two-sample t-test is a rather robust testing procedure. Such optimism may even suggest that normal theory can provide an adequate approximation to the distribution of t_j under H_1.

In a previous Desmatics, Inc. technical report [5], we studied the distribution of t_j when all nonzero factor effects were of equal magnitude. We found that even for relatively small values of N, normal theory provided approximations that agreed very closely with corresponding Monte Carlo estimates. Preliminary indication is that the normal-theory approximations also perform well for an arbitrary set of effects. The approximation states that for fixed N, t_j has approximately a noncentral t-distribution with (N - 2) degrees of freedom and noncentrality parameter

$$\delta_{1j} = N^{\frac{1}{2}}\beta_j/[\tau^2 - \beta_j^2 + \sigma^2]^{\frac{1}{2}} \tag{3.14}$$

To apply this approximation we define

$$\psi_1(\delta) = P\{|T_{N-2}(\delta)| \geq t(N-2;\alpha_1/2)\} \quad , \tag{3.15}$$

where $T_\gamma(\delta)$ denotes a random variable having a noncentral t-distribution with γ degrees of freedom and noncentrality parameter δ, and we let

$$R_{1j} = \begin{cases} 1, & \text{if } H_0:\beta_j = 0 \text{ is rejected in the first stage} \\ 0, & \text{otherwise.} \end{cases}$$

We see immediately that $S = \sum_{j=1}^{K} R_{1j}$ where, as defined previously, S denotes the number of factors carried over to the PB second stage from the RB first stage.

From our definition of R_{1j}, $P(R_{1j} = 1)$ represents the power of the first-stage F-test for detecting the effect of the $j^{\underline{th}}$ factor. The above approxi-

mation to the distribution of t_j has

$$P(R_{1j} = 1) \approx \psi_1(\delta_{1j}) . \tag{3.16}$$

Notice that if $\beta_j = 0$, then $\delta_{1j} = 0$ and $P(R_{1j} = 1) \approx \psi_1(0) = \alpha_1$.

We note further that $E(R_{1j}) = P(R_{1j} = 1)$, so that

$$E(S) = \sum_{j=1}^{K} E(R_{1j})$$

$$\approx \sum_{j=1}^{K} \psi_1(\delta_{1j}) . \tag{3.17}$$

Introducing (3.17) into (2.3) yields

$$E(R) \approx N + 3.5 + \sum_{j=1}^{K} \psi_1(\delta_{1j}) . \tag{3.18}$$

Thus, the expected total number of runs for an RP(N, α_1, α_2) strategy is given approximately by (3.18).

B. SECOND STAGE: PLACKETT - BURMAN DESIGN

Unlike RB experimentation, the testing characteristics of PB designs are fairly well known. For brevity, therefore, we do not rederive these characteristics here but instead apply and state them where needed.

We consider, then, a Plackett-Burman design for the study of S factors in M = B(S + 1) runs, and we again consider testing $H_0: \beta_j = 0$ versus $H_1: \beta_j \neq 0$, assuming of course that factor j is one of the S factors being analyzed, i.e., $R_{1j} = 1$. We define

$$R_{2j} = \begin{cases} 1 & \text{if } H_0: \beta_j = 0 \text{ is rejected in the second stage} \\ 0 & \text{otherwise,} \end{cases}$$

and let

$$\psi_2(s,\delta) = P\{|T_{d(s)}(\delta)| \geq t(d(s);\ \alpha_2/2)\} \tag{3.19}$$

where $d(s) = M - (s+1) = B(s+1) - (s+1)$. The quantity $d(s)$ represents the number of error degrees of freedom in the PB design given that $S = s$ factors are carried over from the RB first stage. We note that if $R_{1j} = 0$, then $R_{2j} = 0$.

Conditional on $S = s$ and on $R_{1j} = 1$, we can show

$$P(R_{2j} = 1 \mid S=s,\ R_{1j} = 1) = \psi_2(s,\ \delta_{2j}) \tag{3.20}$$

where $$\delta_{2j} = [B(s+1)]^{\frac{1}{2}}\ \beta_j/\sigma\ . \tag{3.21}$$

The conditional probability in (3.20) represents the probability that the $j^{\underline{th}}$ factor tests significant from zero in the second stage given that it and $s-1$ other factors test significant from zero in the RB first stage. This is basically the only result we need for the second stage. We re-emphasize that S is a random variable. In subsection C we discuss an approximation to $R(R_{2j} = 1)$, the probability that the $j^{\underline{th}}$ factor is declared important (i.e., $H_0:\beta_j = 0$ is rejected in both stages) by an $RP(N,\alpha_1,\alpha_2)$ strategy.

C. COMBINED STAGES: THE RP(N, α_1, α_2) STRATEGY

To evaluate $P(R_{2j} = 1)$, we observe first that

$$P(R_{2j} = 1) = P(R_{1j} = 1)\ P(R_{2j} = 1 | R_{1j} = 1)\ . \tag{3.22}$$

Since an approximation to $P(R_{1j} = 1)$ is given in (3.16), it remains to consider the conditional probability that $R_{2j} = 1$ given $R_{1j} = 1$.

We define $S_j = S - R_{1j}$. Now,

$$P(R_{2j}=1|R_{1j}=1) = \sum_{s=1}^{K} P(R_{2j}=1,\ S_j = s-1|R_{1j}=1)$$

$$= \sum_{s=1}^{K} P(S_j = s-1|R_{1j}=1)\ P(R_{2j}=1|S_j=s-1,\ R_{1j}=1)$$

$$= \sum_{s=1}^{K} P(S_j = s-1|R_{1j}=1)\ \psi_2(s,\delta_{2j}). \qquad (3.23)$$

To evaluate (3.23) we must somehow approximate the conditional distribution of S_j given $R_{1j}=1$. We lament that this is not a straightforward exercise; the conditional distribution of S_j given $R_{1j}=1$ is extremely complex. Nevertheless, in many screening situations, for moderately sized N the conditional distribution of S_j given $R_{1j}=1$ might be reasonably approximated as the convolution of $(K-1)$ independent Bernoulli random variables having success probabilities $\{\psi_1(\delta_{1m}): m=1, 2, \ldots, j-1, j+1, \ldots, K\}$. Alternatively, following Feller [3], for large N and moderate values of $\lambda_j = \sum_{\substack{m=1 \\ m\neq j}}^{K} \psi_1(\delta_{1m}) = E(S) - \psi_1(\delta_{1j})$, we might reasonably approximate the conditional distribution of S_j given $R_{1j}=1$ by a Poisson distribution having mean λ_j. The Poisson approximation is generally much easier to apply than the Bernoulli convolution approximation, particularly when K is large.

We note that the means of both approximating distributions agree with the conditional mean of S_j given $R_{1j}=1$. The same cannot be said of their variances however. If we define $P_{ij} = P\{R_{1i}=1,\ R_{1j}=1\}$ for any i and j, then

$$V(S_j|R_{1j}=1) = \sum_{i\neq j} P_{ii}(1-P_{ii}) + \sum_{r\neq j} \sum_{\substack{m\neq j \\ r\neq m}} (P_{rr} - P_{rr}P_{mm}). \qquad (3.24)$$

The variance of the Bernoulli convolution approximation is given by $\sum_{i \neq j} P_{ii}(1 - P_{ii})$; the variance of the Poisson approximation is given by $\lambda_j = \sum_{i \neq j} P_{ii}$. Perhaps we could obtain a more refined approximating distribution if we could equate not only the first order but also the second order moments. Unfortunately, we have not yet been able to reasonably approximate P_{rm}, $r \neq m$.

D. MONTE CARLO RESULTS

As a check on the various approximations presented, we conducted three Monte Carlo case studies. The results are summarized in Tables 1, 2, and 3. As can be seen from these tables, the results are extremely encouraging and suggest that the approximations of E(R), $P(R_{1j} = 1)$, and $P(R_{2j} = 1)$, presented herein, are quite reasonable.

As can be noted from the tables, the approximations of $P(R_{2j} = 1)$ based on the Bernoulli convolution and Poisson distribution approaches yield essentially the same results in the first two case studies. Because of this agreement and the complexity of the calculations associated with the Bernoulli convolution for the third case study, we used only the Poisson approximation in that case.

Of the three Monte Carlo case studies, most noteworthy is the third case study, which we feel is more akin to situations encountered in practice than the first and second case studies. In this case, nonzero factor effects vary in magnitude, ranging between σ and 6σ in absolute magitude. Moreover, the majority of the effects are relatively small. Even here the approximations do quite well.

	E(R)	$j \leq 10$		$j \geq 11$	
		$P(R_{1j} = 1)$	$P(R_{2j} = 1)$	$P(R_{1j} = 1)$	$P(R_{2j} = 1)$
$\alpha_1 = .1$					
Approximation	43.25	.475	.452*,.451**	.100	.005*,.005**
Monte Carlo	42.86(.15)	.459(.006)	.432(.007)	.009(.001)	.005(.001)
$\alpha_1 = .3$					
Approximation	67.71	.721	.712*,.712**	.300	.015*,.015**
Monte Carlo	67.54(.23)	.720(.005)	.713(.006)	.299(.001)	.015(.001)
$\alpha_1 = .5$					
Approximation	90.89	.839	.836*,.836**	.500	.025*,.025**
Monte Carlo	90.61(.27)	.838(.005)	.837(.005)	.497(.001)	.024(.001)
$\alpha_1 = .7$					
Approximation	113.65	.915	.914*,.914**	.700	.035*,.035**
Monte Carlo	113.54(.23)	.913(.004)	.913(.004)	.699(.001)	.035(.001)

*Based on Bernoulli convolution approximation.
**Based on Poisson approximation.

Case Study I: K = 120 Factors, N = 24 Runs, $\alpha_1 = .1, .3, .5, .7$, $\alpha_2 = .05$

$$|\beta_j|/\sigma = \begin{cases} 4. & \text{for } j \leq 10 \\ 0. & \text{for } j \geq 11 \end{cases}$$

Table 1: Results of Case Study I. (Value in parentheses represents estimated standard deviation of Monte Carlo estimate.)

	E(R)	$j \leq 3$		$j \geq 4$	
		$P(R_{1j} = 1)$	$P(R_{2j} = 1)$	$P(R_{1j} = 1)$	$P(R_{2j} = 1)$
$\alpha_1 = .1$					
Approximation	23.32	.708	.605*,.605**	.100	.005*,.005**
Monte Carlo	23.37(.13)	.669(.012)	.575(.014)	.103(.002)	.005(.001)
$\alpha_1 = .3$					
Approximation	35.29	.896	.814*,.814**	.300	.015*,.015**
Monte Carlo	35.00(.17)	.885(.008)	.793(.012)	.296(.003)	.015(.001)
$\alpha_1 = .5$					
Approximation	46.86	.952	.891*,.890**	.500	.025*,.025**
Monte Carlo	46.32(.18)	.947(.006)	.889(.009)	.490(.003)	.024(.001)
$\alpha_1 = .7$					
Approximation	58.33	.978	.932*,.927**	.700	.035*,.035**
Monte Carlo	58.46(.17)	.977(.004)	.926(.004)	.699(.003)	.035(.001)

*Based on Bernoulli convolution approximation.
**Based on Poisson approximation.

Case Study II: K = 60 Factors; N = 12 Runs; $\alpha_1 = .1, .3, .5$, and $.7$; $\alpha_2 = .05$

$$|\beta_j|/\sigma = \begin{cases} 2.5 & \text{for } j \leq 3 \\ 0. & \text{for } j \geq 4 \end{cases}$$

Table 2: Results of Case Study II. (Value in parentheses represents estimated standard deviation of Monte Carlo estimate.)

		Monte Carlo			Approximation		
		E(R)	$P(R_{1j}=1)$	$P(R_{2j}=1)$*	E(R)	$P(R_{1j}=1)$	$P(R_{2j}=1)$*
$\alpha_1 = .1$		48.00(.20)			48.33		
$\lvert\beta_j\rvert/\sigma$	No.						
0.	30		.099(.003)	.005(.001)		.100	.005
1.	24		.112(.004)	.072(.003)		.109	.071
2.	18		.123(.004)	.105(.004)		.136	.117
3.	12		.182(.006)	.167(.006)		.180	.165
4.	9		.238(.008)	.226(.007)		.242	.230
5.	5		.318(.011)	.310(.011)		.320	.311
6.	2		.381(.020)	.375(.020)		.412	.406
$\alpha_1 = .3$		69.43(.25)			69.82		
$\lvert\beta_j\rvert/\sigma$	No.						
0.	30		.297(.005)	.015(.001)		.300	.015
1.	24		.315(.005)	.253(.006)		.313	.251
2.	18		.335(.006)	.309(.006)		.351	.322
3.	12		.406(.008)	.390(.008)		.410	.395
4.	9		.489(.009)	.483(.009)		.486	.480
5.	5		.573(.013)	.571(.013)		.573	.571
6.	2		.638(.020)	.637(.020)		.664	.664

*Based on Poisson approximation.

Case Study III: K = 100 Factors α_1 = .1 and .3

N = 30 Runs α_2 = .05

Table 3: Results of Case Study III. (Value in parentheses represents estimated standard deviation of Monte Carlo estimate.)

IV. PRACTICAL CONSIDERATIONS

Although analyzing each factor separately in an RB design via an F-test is a relatively quick and simple testing procedure, it is not necessarily the most powerful method of analyzing data from a random balance design. Presumably, more sophisticated statistical techniques (such as least squares stepwise methods) that analyze more than one factor at a time would provide greater power. Such methods, however, may not be computationally feasible if K is extremely large and may be severely limited if N is very small relative to K. In addition, evaluating the efficacy of many of these methods will be a difficult, if not insurmountable, task. The individual F-test approach, as we have seen, admits to a tractable assessment that may provide a lower bound for the performance characteristics of these alternative, more sophisticated analysis techniques.

The power of the K separate F-tests can, of course, be increased directly by using a larger sample size N or a larger level of significance α_1. However, by increasing N we increase our testing cost by requiring more first-stage runs, and by increasing α_1 we increase not only the probability of first-stage Type I error (i.e., rejecting H_0 when it is true), but also the expected number of factors carried over to the second stage. Increasing α_1, then, will also increase our expected testing cost by requiring, on the average, more second-stage runs.

Similarly, we can increase the power of the second-stage analysis by using a larger PB design or a larger level of significance α_2. Unlike adjusting α_1, however, adjusting α_2 does not affect the expected total number of runs required by an RP strategy. The power of the second-stage analysis might also be increased if some factor effects can be reasonably assumed to

be negligible; in which case we can pool their associated sum of squares into the error sum of squares to obtain a pooled error estimate having more degrees of freedom for error than the unpooled estimate. This is particularly tempting when $d(s) = 1$ where $d(s)$ is defined as in (3.19). Caution, however, must be exercised whenever effects are pooled since pooling has a tendency to diminish the denominator expected mean square of the F-statistic. Nonetheless, because of the relatively large gain in power that accrues for each additional degree of freedom (up to a total of about 5), pooling may be very attractive.

To help determine which effects, if any, might be reasonable to combine in the second-stage analysis, estimated effects can be plotted on normal probability paper. In this technique, due to Daniel [2], small effects should fall approximately along a straight line, while large effects should tend to fall far from the line. This technique by itself can also serve as an alternative method of analysis, although it generally relies heavily on subjective judgement.

Further, in our evaluation of performance in Section III, we assumed for simplicity that all tests of significance in the first and second stages are conducted at the same levels, namely, α_1 and α_2, respectively. In practice, however, differing levels of significance may be used. Larger significance levels could conceivably be used when testing factors anticipated prior to experimentation to have a major effect on the response. This would afford us greater flexibility in regulating the effect of Type I and Type II errors. We note that the results of Section III can be easily modified to allow for distinct levels of significance in the first or second stage. Using different significant levels for different factors implies the incorporation of prior knowledge into a screening strategy. Thus a Bayesian framework

would be called for. However, we will not venture into Bayesian territory in this report.

Finally, we note that the RB first stage can be used alone as a method of factor screening. An immediate advantage of a screening plan based solely on an RB design is that the total number of runs used for screening can be fixed prior to experimentation. A quantitative evaluation of the advantages and disadvantages of using a one-stage versus a two-stage procedure will be addressed in a forthcoming report.

V. SUMMARY

In this report we have endeavored to evaluate the efficacy of the RP screening method. Of course, in order to select an RP strategy, we must consider both the accuracy of factor classification and the number of runs required. Because of the intractable nature of an exact solution, we have developed approximations to (1) the probabiltiy that a given factor is declared important and (2) the expected number of required runs. When compared with the results of three Monte Carlo studies, these approximations fared extremely well.

The results of this report can be used as a practical guide in decisions about the possible use and choice of an RP strategy. In many ways the actual selection of an RP strategy is similar to the process of specifying the sample size for an analysis of variance problem. That is, we must consider trade-offs between Type I error, power, and total number of runs required.

VI. REFERENCES

[1] Budne, T. A., "The Application of Random Balance Designs," Technometrics, Vol. 1, pp. 139-155, 1959.

[2] Daniel, C., "Use of Half-Normal Plots in Interpreting Factorial Two-Level Experiments," Technometrics, Vol. 1, pp. 311-341, 1959.

[3] Feller, W., An Introduction to Probability Theory and Its Applications, Vol. 1, Third Edition, John Wiley & Sons, Inc., New York, NY, 1968.

[4] Kleijnen, J. P. C., "Screening Designs for Poly-Factor Experimentation," Technometrics, Vol. 17, pp. 487-493, 1975.

[5] Mauro, C. A., and Smith, D. E., "Factor Screening in Simulation: Performance of a Two-Stage Random Balance/Plackett-Burman Procedure," Technical Report No. 113-5, Desmatics, Inc., 1981.

[6] Plackett, R. L., and Burman, J. P., "The Design of Optimum Multifactor Experiments," Biometrika, Vol. 33, pp. 305-325, 1946.

[7] Satterthwaite, F. E., "Random Balance Experimentation," Technometrics, Vol. 1, pp. 111-137, 1959.

[8] Smith, D. E., and Mauro, C. A., "Factor Screening in Computer Simulation," Simulation, Vol. 38, pp. 49-54, 1982.

[9] Srivastava, J. N., "Designs for Searching Non-negligible Effects," A Survey of Statistical Design and Linear Models, ed. J. N. Srivastava, North-Holland Publishing Co., New York, NY, 1975.

APPENDIX

If H is a discrete random variable having probability distribution

$$P(H=h) = \begin{cases} \binom{r}{h}^2 \Big/ \binom{2r}{r} & h = 0, 1, 2, \ldots, r \\ 0 & \text{otherwise,} \end{cases} \tag{A.1}$$

where r is any positive integer, we write $H \sim H(r)$. The class of distributions defined in (A.1) is a symmetric subfamily of the hypergeometric family of distributions. If $H \sim H(r)$, then $E(H) = r/2$ and $V(H) = (r/2)^2/(2r-1)$.

We define $\underline{f}_i = (\underline{x}_i + \underline{1})/2$ where $\underline{x}_i$ is the $i^{\underline{th}}$ column vector of an RB design matrix and note that, for $i \neq j$, $\underline{f}_j'\underline{f}_i \sim H(N/2)$. It follows that, for $i \neq j$, $(\underline{x}_j'\underline{x}_i + N)/4 \sim H(N/2)$. Hence, for $i \neq j$, $E(\underline{x}_j'\underline{x}_i) = 0$ and $V(\underline{x}_j'\underline{x}_i) = N^2/(N-1)$.

In regard to the distribution of $\underline{x}_j'\underline{\varepsilon}$, we observe that $\underline{x}_j'\underline{\varepsilon}|\underline{x}_j \sim N(0, N\sigma^2)$ since $\underline{x}_j'\underline{x}_j = N$. Because the conditional distribution of $\underline{x}_j'\underline{\varepsilon}$ given $\underline{x}_j$ is the same for any realization of $\underline{x}_j$, the result is therefore true unconditionally. Thus, $E(\underline{x}_j'\underline{\varepsilon}) = 0$ and $V(\underline{x}_j'\underline{\varepsilon}) = N\sigma^2$.

To find the covariance between $\hat{\beta}_i$ and $\hat{\beta}_j$, for $i \neq j$, we have

$$\operatorname{cov}(\hat{\beta}_i, \hat{\beta}_j) = E(\hat{\beta}_i\hat{\beta}_j) - \beta_i\beta_j \ . \tag{A.2}$$

We write $N\hat{\beta}_m = \gamma_m + \lambda_m$ where $\lambda_m = \underline{x}_m'\underline{\varepsilon}$ and $\gamma_m = \sum_{q=1}^{K} \beta_q \underline{x}_m'\underline{x}_q$. Thus,

$$N^2E(\hat{\beta}_i\hat{\beta}_j) = E(\gamma_i\gamma_j) + E(\gamma_i\lambda_j) + E(\gamma_j\lambda_i) + E(\lambda_i\lambda_j). \tag{A.3}$$

It is easy to show that $E(\gamma_i\lambda_j) = E(\gamma_j\lambda_i) = E(\lambda_i\lambda_j) = 0$, so that

$$\operatorname{cov}(\hat{\beta}_i, \hat{\beta}_j) = E(\gamma_i\gamma_j)/N^2 - \beta_i\beta_j \ . \tag{A.4}$$

Expanding $E(\gamma_i\gamma_j)$, we have

$$E(\gamma_i\gamma_j) = \sum_{r=1}^{K}\sum_{q=1}^{K} \beta_r\beta_q\, E(\underline{x}_i'\underline{x}_r\underline{x}_j'\underline{x}_q) \quad . \tag{A.5}$$

It is not difficult to verify that, for $i \neq j$, $E(\underline{x}_i'\underline{x}_r\underline{x}_j'\underline{x}_q)$ is zero unless $r=i$, $q=j$ or $r=j$, $q=i$. When $r=i$ and $q=j$, $E(\underline{x}_i'\underline{x}_r\underline{x}_j'\underline{x}_q) = E[(\underline{x}_i'\underline{x}_i)^2] = N^2$; when $r=j$ and $q=i$, $E(\underline{x}_i'\underline{x}_r\underline{x}_j'\underline{x}_q) = V(\underline{x}_j'\underline{x}_i) = N^2/(N-1)$.

Substitution into (A.5) yields

$$E(\gamma_i\gamma_j) = N^3\beta_i\beta_j/(N-1). \tag{A.6}$$

Finally, introducing (A.6) into (A.4), we get the desired result

$$\mathrm{cov}(\hat{\beta}_i,\hat{\beta}_j) = \beta_i\beta_j/(N-1) \quad . \tag{A.7}$$

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: 113-6	2. GOVT ACCESSION NO.: AD-A117529
3. RECIPIENT'S CATALOG NUMBER:	
4. TITLE (and Subtitle): FACTOR SCREENING IN SIMULATION: EVALUATION OF A RANDOM BALANCE/PLACKETT-BURMAN STRATEGY	5. TYPE OF REPORT & PERIOD COVERED: Technical Report
	6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Carl A. Mauro and Dennis E. Smith	8. CONTRACT OR GRANT NUMBER(s): N00014-79-C-0650
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Desmatics, Inc. P. O. Box 618 State College, PA 16801	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR 277-291
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research Arlington, VA 22217	12. REPORT DATE: July 1982
	13. NUMBER OF PAGES: 24
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this report is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Factor Screening
Computer Simulation
Random Balance Design
Plackett-Burman Design

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Large, complex simulation models usually must be studied by performing a number of simulation runs. It is often to our advantage to invest a relatively small number of runs in a preliminary experiment aimed at screening out the most important factors. The major experimental effort and expense can then be focused on these key factors. In this report we evaluate the performance of a two-stage screening strategy that is based on a combination of random balance and Plackett-Burman designs.

DD FORM 1 JAN 7 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)